

EXHIBIT H

From: [Kathleen Hartnett](#)
To: [Reed Forbush](#); [Lauter, Judd](#); [Jesse Panuccio](#); [Maxwell Pritt](#); [Holden Benon](#); [Christopher Young](#); [Aaron Cera](#); [Cadio Zirpoli](#); [Joe Saveri](#); [Margaux Poueymirou](#); [Ashleigh Jensen](#); [Rya Fishman](#); [Matthew Butterick](#); [Nada Djordjevic](#); [James Ulwick](#); [Bryan L. Clobes](#); [Mohammed Rathur](#); [Amy Keller](#); [David Straite](#); [Ruby Ponce](#); [Alexander Sweatman](#); [Heaven Haile](#); [Llama BSF](#); [Josh Schiller](#); [David Boies](#)
Cc: [Ghazarian, Colette A](#); [Poppell, Cole A](#); [Dunning, Angela L.](#); [Ghajar, Bobby A.](#); [Alvarez, Jessica](#); [Weinstein, Mark](#); [Stameshkin, Liz](#); [z/Meta-Kadrey](#)
Subject: RE: Kadrey v. Meta - October 16 Meet and Confer
Date: Monday, October 21, 2024 8:26:00 PM

CAUTION: External email. Please do not respond to or click on links/attachments unless you recognize the sender.

Reed,

We have reviewed your responses below. Reserving all rights, including whether the issues below reflect impasses and/or have been the subject of sufficient conferral, we do not object to you contacting Judge Hixson's chambers to request a telephonic conference. We too hope to resolve issues where we can prior to any call with Judge Hixson or motions practice. We also reserve our right to add proposed issues to discuss with Judge Hixson regarding Plaintiffs' discovery responses.

Thanks,
Kathleen

Kathleen R. Hartnett
Cooley LLP
3 Embarcadero Center, 20th Floor
San Francisco, CA 94111-4004
+1 415 693 2071 office
+1 415 693 2222 fax
+1 202 368 5916 mobile
khartnett@cooley.com
www.cooley.com/people/kathleen-hartnett
Pronouns: she, her, hers

From: Reed Forbush <rforbush@BSFLLP.com>
Sent: Monday, October 21, 2024 12:35 PM
To: Lauter, Judd <jlauter@cooley.com>; Hartnett, Kathleen <khartnett@cooley.com>; Jesse Panuccio <jpanuccio@BSFLLP.com>; Maxwell Pritt <mpritt@BSFLLP.com>; Holden Benon <hbenon@saverilawfirm.com>; Christopher Young <cyoung@saverilawfirm.com>; Aaron Cera <aCera@saverilawfirm.com>; Cadio Zirpoli <czirpoli@saverilawfirm.com>; Joe Saveri <jsaveri@saverilawfirm.com>; Margaux Poueymirou <mpoueymirou@saverilawfirm.com>; Ashleigh Jensen <ajensen@saverilawfirm.com>; Rya Fishman <rfishman@saverilawfirm.com>; Matthew Butterick <mb@buttericklaw.com>; Nada Djordjevic <ndjordjevic@dicellolevitt.com>; James Ulwick <Julwick@dicellolevitt.com>; Bryan L. Clobes <BClobes@caffertyclobes.com>; Mohammed Rathur <MRathur@caffertyclobes.com>; Amy Keller <akeller@dicellolevitt.com>; David Straite <dstraite@dicellolevitt.com>; Ruby Ponce <rponce@saverilawfirm.com>; Alexander Sweatman <ASweatman@caffertyclobes.com>; Heaven Haile <hhaile@saverilawfirm.com>; Llama BSF

<Llama_BSF@bsfllp.com>; Josh Schiller <JiSchiller@BSFLLP.com>; David Boies
<DBoies@BSFLLP.com>

Cc: Ghazarian, Colette A <cghazarian@cooley.com>; Poppell, Cole A <CPoppell@cooley.com>;
Dunning, Angela L. <adunning@cgsh.com>; Ghajar, Bobby A. <bghajar@cooley.com>; Alvarez,
Jessica <jalvarezlopez@cooley.com>; Weinstein, Mark <mweinstein@cooley.com>; Stameshkin, Liz
<lstameshkin@cooley.com>; z/Meta-Kadrey <zmetakadrey@cooley.com>

Subject: RE: Kadrey v. Meta - October 16 Meet and Confer

[External]

Counsel:

We reply to your response of earlier this morning to our correspondence on the parties' 10/16 meet and confer, and specifically to the portion below containing your individual responses on the various issues discussed. Plaintiffs' responses to those responses are interlineated below in bold and red font.

We dispute your comments on timing and as to the presence of any "new issues." We note generally that on many subjects for which you stated that Meta would return with supplemented responses, you have provided little in the form of such supplemented responses and have not made any promises to produce any missing responses by specific dates.

Given the short timeframe, we must preserve our rights (and avoid prejudice from any further delay in resolving discovery disputes) by teeing up such issues for the Court's consideration. Please let us know if you will not object to our contacting chambers to arrange a telephonic conference, as provided in Judge Hixson's standing order. Otherwise, we will do so in the status report that Judge Hixson has ordered to be filed tomorrow by noon. Naturally, if any issues can be resolved before the court hears or rules on them, then we can withdraw them.

Best,
Reed

I. **"Follow-Up Issues"**

1. RFP No. 64: Meta will consider proposing an agreeable limitation and searching/producing responsive documents.

[Meta response: This request seeks documents showing "each instance within the last three years where You have licensed copyrighted works for Meta's commercial use," without any limitation whatsoever, including any tie to Meta's training of the Meta Language Models. As we discussed on the call, Meta objected to this request 6 months ago and Plaintiffs never raised it since, and it is unduly burdensome and not proportional to the needs of the case. In response to your latest set of RFPs, Meta will be making a production regarding the licensing of copyrighted textual works for Meta's use in training the Meta Language Models, which reflects the scope of potentially relevant materials in this case. We believe that this production will be sufficient and appropriate to respond to RFP No. 64 and the latest requests. We trust that this issue is now resolved.]

The date of Meta's objection is irrelevant. Meta also continues to use its defined terms "Meta Language Models," which as discussed does not include all Llama models, including Llama 4 even if the RFP was tied to Llama, which it's not. Also, "Making a production" is vague, as is

“reflects the scope.” We cannot rely on these statements, so do not agree that this issue is resolved. The RFP is clear and plainly relevant. The parties therefore remain at an impasse.

2. LibGen Training Data (RFP Nos. 1-3 & 7, ROG No. 1): Meta declined to produce all datasets with copyrighted material that it downloaded but will consider whether it would identify all iterations of downloaded datasets with copyrighted material (including the Bashlykov download of LibGen).

[Meta Response: Since the beginning of this case, Meta has been clear that it would produce the training data alleged to contain the Plaintiffs’ works and has done so, along with significant additional training data. The training data produced by Meta constitutes many terabytes of data, and Plaintiffs’ counsel has complained that this volume of data is too large for them to handle. In any event, Meta had understood that the parties had reached a compromise to limit production of datasets to those datasets that are alleged to contain Plaintiffs’ works. It is not proportional to the needs of the case, nor pertinent to adjudication of the fair use defense, to produce copies of all pre-training and fine tuning datasets, but Meta is willing to consider producing additional datasets, if any, that contain books. In addition, Meta will confirm whether the Libgen dataset that has been produced is the current and most comprehensive version of that dataset. However, Meta will not identify and produce duplicative copies of each training dataset, which, among other things, would impose a highly disproportionate burden in light of Plaintiffs’ needs.]

As we have explained, there is no issue more relevant than the copying of the infringing datasets, and the intentions behind and use of each infringing dataset is especially pertinent to fair use. All such datasets must be produced to show the scope of Meta’s infringement. Meta has not made a showing of any burden or proposed a reasonable alternative that would still demonstrate the scope of infringement, particularly in light of our proposal to provide information re: other downloads rather than the downloads themselves. We thus remain at an impasse.

3. GitHub-Hosted Communications and Source-Code Repositories: Meta will investigate gaps in data (e.g., time periods during which there were no pull requests) as identified in declaration of Joannathan Krein.

[Meta Response: We confirm that we are actively identifying any alleged gaps in the pull requests and will produce any located pull requests for the requested time periods. We are also investigating whether additional source code should be produced in response to Plaintiffs’ most recent set of discovery requests (as Plaintiffs’ initial requests did not call for source code). We thus trust that this issue is now resolved.]

Plaintiffs thank Meta for agreeing to produce all responsive pull requests. Please provide a date certain upon which we will receive the same. Meta has now had two weeks since receiving our letter regarding this deficiency and our request for complete source code for longer, so saying you’ll investigate further is insufficient at this late stage. Moreover, all of source code is clearly relevant and responsive to the RFPs at issue. We thus remain at an impasse.

4. Llama Filters or Training Data (RFP Nos. 45, 64, 74, 76, and 77): Meta will consider providing screenshot(s) of the filtering tool that Meta’s counsel described on the call and represented was not collectible.

[Meta Response: None of the cited RFPs address any alleged tools for filtering data. Moreover, the “tool” you are referring to is merely a tool used to visualize data. Because (among other things)

there is no identified RFP to which this tool would be responsive, we do not agree there is anything else to provide in response to this request and we trust that this issue is now resolved.]

As stated during our meet and confer, any tools—whether or not you want to call them something else, like processes, programs, methods, etc.—for filtering, especially if they can filter copyrighted material, are plainly relevant and responsive. You confirmed there was a “tool” utilized by Meta and we compromised by offering to accept screenshots of it if it couldn’t otherwise be produced. Moreover, forcing Plaintiffs to issue a new (duplicative and/or overlapping) RFP is necessarily wasteful of time and resources. We thus remain at an impasse.

5. Valuation and Financials: Meta will confirm the scope of data/documents already searched and produced (and whether the documents pertain to Llama 4—see general impasse regarding Llama 4.)

[Meta Response: Meta has already produced financial-related materials responsive to Plaintiffs’ document requests. Meta will be producing additional financial-related documents responsive to Plaintiffs’ latest set of RFPs in due course to the extent any such documents exist. Meta has not withheld responsive financial related documents on the basis that they relate only to Llama 4. We thus trust that this issue is now resolved.]

It is not clear whether Meta searched for and produced all responsive information, yet it is clear Meta is withholding documents on account of its assertion that it would only be responsive to the “latest set” of RFPs and not responsive to prior RFPs. We thus remain at an impasse.

6. Identifying Responsiveness of Document Productions: Meta said it’ll consider including in its document production letters an indication of the content of the productions and to what RFPs they’re responsive.

[Meta Response: We discussed notifying Plaintiffs when a production addresses a production or formatting issue from a prior production. We are not aware of, and you did not provide, any authority supporting an obligation for a party to provide descriptions of the content of its document productions and to tie those productions to specific RFPs. To resolve this issue, going forward we will provide a description of our productions when we make them. We thus trust that this issue is now resolved.]

We appreciate your agreement to provide descriptions of productions when made. Please let us know if that will include identifying RFPs to which the documents contained in a production are responsive, without prejudice to identifying every possible RFPs to which any particular document may be responsive as we recognize documents can be responsive to multiple RFPs and RFPs can overlap. Plaintiffs would agree to do the same. If acceptable, then this issue is resolved.

7. Identification of Search Terms: Meta declined to identify search terms and hit counts unless it was a mutual exchange and said it would get back to us on when it could exchange search terms and hit counts. Meta said it would identify all non-custodial and custodial sources searched but not all sources with potentially relevant information (see Identification of Search Terms entry under Impasse section).

[Meta Response: As we stated, we are agreeable to engaging with Plaintiffs in a mutual exchange

of search terms and hit counts as contemplated by the parties' agreed-upon ESI order. (Doc No. 101, para. 7). This disclosure would also identify which non-custodial and custodial sources of information that were searched with these terms. We do not believe there are any other types of data sources with potentially relevant information that were not searched for and/or collected from, nor is that required by the ESI order. We trust that this issue is now resolved.]

We asked you to let us know when you can provide search terms and hit counts for the mutual exchange you demanded, and you still have not done so. As we explained, a party must produce all known responsive information and documents in response to discovery requests after a reasonable, good faith investigation for responsive information. That investigation is not limited only to document custodians. See, e.g., *Strategic Partners, Inc. v. FIGS, Inc.*, No. 19-CV-2286-GW (KSX), 2020 WL 2527056, at *6-7 (C.D. Cal. Feb. 6, 2020) (defendant "impermissibly truncated the scope of its search for responsive information" by limiting its custodians because the "reasonable inquiry" required under Rule 26 must include, "at a minimum, a reasonable procedure to distribute discovery requests to *all* employees and agents of the [party] *potentially possessing* responsive information") (emphasis in original). Meta may not conceal information and refuse to produce known responsive information on the basis that the ESI Order sets forth how documents may be collected from document custodians. See, e.g., Dkt. 101, § 7(d) ("Specific, non-duplicative ESI that is identified by a party as responsive to a discovery request shall not be withheld from review or production solely on the grounds that it was not identified by (or is subject to an exclusion set forth in) the protocols described in, or developed in accordance with, this Order."). We thus remain at an impasse.

8. RFP No. 59: Meta said it would consider proposing a limitation or definition for "fictional works."

[Meta Response: We asked Plaintiffs to explain what they mean by documents about the output of "fictional works." We need to understand what Plaintiffs are seeking here in order to be able to respond to this request.]

Plaintiffs continue to be puzzled by Meta's contention that it doesn't understand what "fictional works" are. We explained "fictional" has the same meaning as defined in dictionaries, and "works" refers to copyrighted works. This request cannot arguably be construed by Meta as too vague so as to not warrant a response. We thus remain at an impasse.

9. Interrogatory No. 3: Meta agreed to supplement its response.

[Meta Response: Yes, Meta will supplement this response. We trust that this issue is now resolved.]

Thank you. Please let us know by what date you will supplement your response to this interrogatory.

10. Interrogatory No. 4: Meta said it would consider proposing an agreeable construction of this interrogatory.

[Meta Response: We do not know what you mean by "agreeable construction." We said that we would take another look at the interrogatory and consider supplementation. We will supplement our response. We trust that this issue is now resolved.]

Thank you. Please let us know by what date you will supplement your response to this

interrogatory.

11. Interrogatory No. 5: Meta said it would investigate Plaintiffs' description of inchoate agreements as negotiations for licenses, deals in process, or deals that fell apart.

[Meta Response: As an initial matter, we do not understand what you mean by "inchoate agreement," as by definition an agreement is something that is not inchoate. In any event, we will search for and produce any documents about the acquisition or licensing of training data in response to more recently served RFPs. We will also supplement our response to this interrogatory accordingly. We trust that this issue is now resolved.]

We similarly are puzzled by Meta's continued failure to understand what an "inchoate agreement" is. But if Meta prefers the term "draft and/or unexecuted agreement," it may construe the term that way. Absent agreement to supplement your response, we remain at an impasse.

12. Interrogatory Nos. 13 and 14: Meta agreed to answer the sub-parts in Rog Nos. 13 and 14.

[Meta Response: While we maintain and do not waive any objections to the fact that these interrogatories and others previously served by Plaintiffs contain multiple subparts, Meta will supplement its responses. We trust that this issue is now resolved.]

Thank you. Please let us know by what date you will supplement your response to this interrogatory.

13. Data Formats: Meta confirmed some garbled chats are from WhatsApp and represented its production of Workplace Chats are in the format available. Meta said it would consider providing file types (but did not agree to re-produce chats in any other formats, see entry in Impasses section).

[Meta Response: We are working to reproduce the discrete number of WhatsApp chats that had a processing issue. With respect to Workplace chats, you have not identified any legitimate issues with the format of those materials. All text is legible, including emojis and links, in the document you cited. To your question, the format in which they were archived, searched, and collected has a file extension of eMail. This reproduction and information about file extension should address any issues that exist re "data format." We trust that this issue is now resolved.]

We do not understand what you mean by "discrete number"—we asked Meta to reproduce all responsive WhatsApp chats in a reasonably usable format similar to how chats are viewed by users. While reserving our rights regarding the dispute on whether Meta has produced WhatsApp and the Workplace chats in a reasonably usable format, we currently are attempting to ascertain whether we might be able to process the files Meta provided in an adequate format and thus consider this issue resolved for the moment unless and until we learn that they cannot be processed on our end for sufficient legibility, provided that Meta will agree not to object to showing the Jury this evidence once so processed.

II. **"Potentially Disputed Issues"**

1. Production of Incorrectly Imaged Documents: Meta said it was aiming for the end of October to re-produce documents incorrectly imaged and/or formatted that Judge Hixson ordered on 10/1 it must produce. Plaintiffs will ask Judge Hixson to impose an earlier deadline given the

deposition schedule.

[Meta Response: Meta has already addressed and reproduced 1,335 documents with imaging issues, which is a majority of the documents at issue. There are no more than 1,100 documents remaining with potential imaging issues. We anticipate completing our review and reproduction by no later than the end of the month and will endeavor to do so sooner. We trust that this issue is now resolved.]

Please let us know if Meta agrees to complete this re-production by this Friday. If so, we agree this issue is resolved.

2. General Scope of Custodial Files Searched:

Meta declined to search custodial files outside of the 10 custodians it originally selected and the additional five ordered by Judge Hixson. Meta also represented its search for responsive documents was limited to non-custodial sources and custodial sources of only those 10 custodians (and now the additional five). Meta also identified two for which it searched custodial files beyond its initial 10 custodians but declined to identify whether it conducted a similar search for any other requests. This dispute concerns all RFPs, including Communications with Business Partners and Communications with Shadow Libraries.

[Meta Response: In compliance with the parties' agreed-upon ESI Order, Meta conducted a comprehensive search for documents responsive to Plaintiffs' requests and conducted custodial documents for the designated custodians. The ESI Order limited custodial ESI searches to only 10 custodians, and Judge Hixson's October 4 order added 5 more custodians. You have identified no authority for your demand that Meta conduct boundless custodial searches beyond the 15 custodians that have already been identified. On our meet and confer, we discussed specific requests regarding, for example, training data discussions with certain companies (RFP 29-34) or archives of information on the internet (RFP Nos. 7-12). We are willing to conduct some reasonable follow up searches for these discrete categories.]

Again, as we explained, Meta must produce all known responsive information and documents in response to discovery requests after a reasonable, good faith investigation for responsive information. That investigation is not limited to non-custodial data sources and only a handful of document custodians. See, e.g., *Strategic Partners, Inc. v. FIGS, Inc.*, No. 19-CV-2286-GW (KSX), 2020 WL 2527056, at *6-7 (C.D. Cal. Feb. 6, 2020) (defendant "impermissibly truncated the scope of its search for responsive information" by limiting its custodians because the "reasonable inquiry" required under Rule 26 must include, "at a minimum, a reasonable procedure to distribute discovery requests to *all* employees and agents of the [party] *potentially possessing* responsive information") (emphasis in original). Meta may not refuse to search for responsive information held by other individuals (which, as explained, is not the same as making every employee a document custodian) and conceal responsive information on the basis that the ESI Order sets forth how documents may be collected from document custodians. See, e.g., Dkt. 101, § 7(d) ("Specific, non-duplicative ESI that is identified by a party as responsive to a discovery request shall not be withheld from review or production solely on the grounds that it was not identified by (or is subject to an exclusion set forth in) the protocols described in, or developed in accordance with, this Order."). We thus remain at an impasse.

Meta declined to add any additional custodians identified by Plaintiffs.

[Meta Response: We have already produced documents for the 10 custodians authorized by the ESI

order and are in the process of doing so for 5 additional custodians allowed for by Judge Hixson's recent order. Plaintiffs have identified no reason for doubling (and now in your latest email, tripling) the number of custodians this late in the discovery process. Plaintiffs also have failed to follow the required procedures set forth in the ESI order for seeking additional custodians.]

The relevance of the additional custodians we have proposed is obvious, including those who will be deposed and any who will serve as 30(b)(6) witnesses, and Meta's refusal to search for produce relevant documents from non-custodians further makes it critical that these custodians be added for searches for responsive documents. The parties thus remain at an impasse.

3. All Llama Models and All Datasets: Meta represented that despite its written objections it has included all Llama models in its document and data searches and productions except as to Llama 4, and it declined to search for and produce responsive information and data in connection with Llama 4. Meta also represented that despite its written objections it has not limited its search for and production of datasets to Book3, but it declined to identify all datasets with copyrighted works that it copied or how many times those datasets, in full or in part, were downloaded. Meta also represented it believed only datasets relating to training was relevant and declined to produce other datasets. This dispute concerns Plaintiffs' RFP Nos. 64, 67, and 49, GitHub-Hosted Communications and Source-Code Repositories, documents and interrogatories related to Valuation and Financials.

[Meta Response: This issue relates to Follow Up Item #2 above, and we refer you to our response above. Meta has stated that it would produce the training data alleged to contain the Plaintiffs' works and has done so. That training data constitutes multiple terabytes of data, which Plaintiffs' counsel has complained is too large for them to handle. It is not proportional to the needs of the case, nor pertinent to adjudication of the fair use defense, to produce any duplicative copies of enormous data sets. With respect to Llama 4, that model is being developed and not expected to be released until next year, after the close of fact discovery in this case. While we are willing to work with Plaintiff to identify reasonable discovery to address the fair use issues as it relates to Llama 4, we do not think it is proportional to the needs of the case nor practical to comprehensively address another, unreleased and unfinalized model in the short time period the parties have to conduct the remaining discovery.]

As we've explained and you continue to ignore, we do not seek only "training" datasets, but all datasets with copyrighted material that Meta uses at any stage, including pre- and post-training. We also continue to disagree that Llama 4's later launch date has any relevance to this dispute or is a proper basis to refuse to produce responsive information. We thus remain at an impasse.

4. RFP No. 45: Meta declined to expand its search beyond use of "tools" for identifying licensed material to include documents or materials regarding licensing and communications with book publishers.

[Meta Response: While we maintain our position that there are no "tools" responsive to your request, as discussed above regarding Follow Up Item #1, Meta will be producing documents about acquisition or licensing of training data in response to more recently served RFPs. We consider this issue resolved.]

As stated during our meet and confer, any tools—whether or not you want to call them something else, like processes, programs, methods, etc.—for identifying licensed material is plainly relevant responsive. Referring to later RFPs is not responsive. We thus remain at an

impasse.

5. Production of Hyperlinks: Meta declined to make a supplemental production of all hyperlinked documents.

[Meta Response: As the ESI order makes clear, hyperlinked documents are not akin to attachments (ESI Order, Doc. No. 101 at 17). Meta has been responding to your requests for additional hyperlinked documents; has been producing the hyperlinked documents when they are responsive; and has noted that in many instances you already had the document you were asking us to produce. We expect to complete our production of any remaining responsive hyperlinked documents that you have requested in about a week. We trust that this issue is now resolved.]

The ESI order does not suggest hyperlinked documents must not produced, only that “[l]inks within a document are not considered attachments.” Because we now know that Meta often uses hyperlinks in lieu of attachments, we suggested amending the ESI Order. Meta refused. Absent agreement by Meta to, at a minimum, review and produce all hyperlinked documents within responsive documents that also are relevant and/or responsive, then we remain at an impasse.

6. Addition of Timothy Dettmers as Custodian: Meta declined to add Mr. Dettmers as a custodian.

[Meta Response: We do not agree that Mr. Dettmers should be added as a custodian. To begin, the ESI order has a protocol for adding custodians and Plaintiffs have not followed that protocol here. Moreover, as we have repeatedly explained, and as Plaintiffs are well aware, Mr. Dettmers was a part-time employee at Meta prior to the company’s development of the Llama models and was uninvolved in the development of those models. Mr. Dettmers’ documents, including those that Judge Hixson determined were privileged, have no bearing on the issues in dispute. Moreover, you have known about Mr. Dettmers since at least December 2023 and provide no reason why Plaintiffs’ suddenly need his ESI now, particularly after Judge Hixson’s adjudication of the privilege dispute.]

We do not agree the ESI protocol requires any additional steps to be undertaken, and to the extent Meta insists on adhering to the time frames set forth therein for considering whether to add a document custodian, we will raise that issue with Judge Hixson as it is unworkable. We also are not required to accept your mere assertion that Mr. Dettmers’ documents “have no bearing on the issues in dispute,” particularly when you refuse to search his files. Further, as we explained, the adjudication of privilege and waiver with respect to a single document is irrelevant to the issue of searching his files for responsive documents. We thus remain at an impasse.

7. Identification of Repositories Not Searched: Meta declined to identify repositories of potentially relevant documents that were not searched. This dispute is relevant to all of RFP Nos. 1-12 and 36-38.

[Meta Response: We do not understand what you are referring to . Meta searched for and/or collected responsive documents in the data sources within Meta’s possession, custody or control that would have potentially responsive or relevant documents, including both non-custodial files and custodial files (see response to Potentially Disputed Issue #2 above). Meta is not aware of any types of data sources with potentially responsive or relevant documents that were not searched. We will provide that information when the parties exchange ESI search terms and hit counts (see Follow Up Issue #7 above).]

Meta represented that it did not search several non-custodial data repositories such as the Hive despite their frequent mention in documents Meta produced (see examples in M&C deficiency letters) and made vague statements about the fact that data within these non-custodial sources may not be searchable or available without explanation. We thus remain at an impasse.

8. Identification of Noncustodial and Custodial Sources: Meta declined to identify all sources that potentially might have responsive information even if not searched, and confirmed its investigation of custodial files to search consisted of asking the custodians in a general way what modes of communication they use to communicate about work (without specifically asking about alternative modes like WhatsApp, Discord, Signal, Telegram, etc.). Meta was not able to explain the production of Al-Dahle WhatsApp messages on the night before his deposition.

[Meta Response: Your statements do not reflect what the parties discussed. Meta searched for and/or collected responsive documents in the data sources within Meta's possession, custody or control that would have potentially responsive or relevant documents, including both non-custodial files and custodial files (see response to Potentially Disputed Issue #2 above). For each of the ESI custodians, Meta has made a reasonable inquiry into which modes of communication they use for work, including by asking about particular modes by name. We will identify the data sources (including communication platforms) that were searched when the parties mutually exchange search terms and hit counts.

During the meet and confer, Meta represented that it relied on the representations of individual custodians as to which devices they used for work and did not further inquire as to what personal devices were owned by the custodians and/or what kind of relevant content might be contained on the same. Nor did Meta explain how and why it obtained WhatsApp chats for one witness (on the eve of his deposition) but no other WhatsApp chats have been provided. Meta further refused to conduct a reasonable investigation into the various methods of communication commonly used as requested in our letters. We thus remain at an impasse.

9. Searching Manifold, GTT, Hive: Meta generally refused to perform searches in Manifold, GTT or Hive on the purported grounds that purportedly they cannot be searched and, in any event, would only contain duplicative data. Meta agreed to consider identifying (not producing) the iterations of copies of training datasets with copyrighted material or books within their PCC.

[Meta Response: This item involves storage locations that may include training data, an issue that was previously addressed in Follow Up Item #2 above. Meta has stated that it would produce the training data alleged to contain the Plaintiffs' works and has done so. That training data constitutes multiple terabytes of data, which Plaintiffs' counsel has complained is too large for them to handle. It is not proportional to the needs of the case, nor pertinent to adjudication of the fair use defense, to produce any duplicative copies of enormous data sets.]

As stated above, we remain at an impasse on this request given Meta's refusal to search known sources of responsive data and limitation to "training" datasets and singles copies thereof.

10. RFP No. 54: Meta declined to search for documents on the grounds that the "factual premise" of the RFP is purportedly wrong, even if there had been a decision at one point not to market Llama with a user interface.

[Meta Response: RFP No. 54 asks for documents about a decision "not" to create an interface for

users for the Meta Language Models. But Meta did create an interface for users for the Meta Language Models – Meta AI – and Meta has produced documents about that interface. This document request therefore is based on an inaccurate premise and, in light of that inaccuracy, you have not articulated any reason it would be relevant or proportional to Plaintiffs’ needs.]

Plaintiffs understand that Meta initially developed Llama without an end user interface (ostensibly to be white labeled for use by others). But even if that understanding is wrong (and Meta has yet to actually demonstrate it is), documents about Meta’s planning regarding an end-user interface is still responsive and relevant to the Fair Use inquiry. Forcing Plaintiffs to issue a new RFP without the word “not” is wasteful and would disregard judicial economy. Absent Meta’s agreement to search for and produce all responsive documents concerning its decision(s) to create and/or not create a user interface for Llama models, we remain at an impasse.

11. Interrogatory No. 15: Meta declined to answer all sub-parts of this interrogatory.

[Meta Response: Reserving all rights and objections, Meta intends to supplement its response to this interrogatory. We trust that this issue is now resolved.]

Absent agreement by Meta to supplement its responses to all subparts of this interrogatory, we remain at an impasse. Please let us know if, contrary to your statement otherwise on the meet and confer call, Meta will respond substantively to all subparts of this interrogatory.

12. Additional Interrogatories Requested: Meta declined to agree to Plaintiffs’ request to expand the numbers of interrogatories the parties may serve and answer the interrogatories Plaintiffs served on October 10.

[Meta Response: Meta is willing to agree to a small increase in the number of interrogatories, up to 5 more than the 25 the parties had previously agreed to, provided the increase is mutual for both sides. We trust with this compromise that this issue is now resolved.]

Given the complexity of this case and the clearly relevant nature of Plaintiffs’ recent interrogatories, which you have had for weeks now, we believe the number of interrogatories we requested is appropriate and will help narrow the issues in this case. Further, we disagree that Plaintiffs have exceeded their number of interrogatories—we are entitled to 25 additional interrogatories now that Farnsworth has been consolidated. In any event, absent Meta’s agreement to withdraw its objections to the number of interrogatories served, we remain at an impasse.

13. Data Formats: Meta declined to reproduce in a legible format the Workplace Chat and WhatsApp messages that Plaintiffs explained contain garbled text or was not produced in a reasonably usable format.

[Meta Response: This complaint is duplicative of Follow Up Item #13 above. We have addressed these issues there.]

See Plaintiffs’ reply above as to the same.

14. Relevant Time Period: Meta declined to extend its searches to the beginning of the class period.

[Meta Response: Since the beginning of discovery in this case, Meta has clearly stated that its collection and production of documents would be based on the January 1, 2022 date that it included in its responses and objections. This is a conservative date because Meta did not even begin work on the Meta Language Models at issue in this case until late 2022. Plaintiffs never raised any issue with this time frame until your October 9 letter. We have no reason to believe that there would be relevant responsive documents at an earlier date than 2022 regarding the Llama models.

Plaintiffs cannot rely on Meta's uncorroborated and speculative claim that there are no responsive materials prior to January 1, 2022. In any event, if that were true, then Meta would not have anything to produce so there is no reason not to search that time frame as well. Because the claims period is what it is, and documents and information from within that period that are responsive must be produced and provided, we remain at an impasse. That said, if Meta is willing to provide a declaration stating the precise date(s) it began working on Llama models and that there are no responsive documents earlier than January 1, 2022, we will consider it.

15. Fair Use Interrogatory (No. 1, Set 2): Meta declined to amend its interrogatory to remove its reference to unidentified documents that it contends support its Fair Use defense and to identify the them.

[Meta Response: Your statement does not reflect the parties' discussion. We told you that Meta will supplement its response to this interrogatory, including identifying documents where appropriate. We trust that this issue is now resolved.]

Thank you, but we also asked Meta to withdraw its incorporation of unidentified discovery materials. Absent Meta's agreement to do so (and by a date certain soon), we remain at an impasse.

16. Advice of Counsel Interrogatory (No. 2, Set 2): Meta would not agree to amend its response to state without ambiguity and equivocation whether it will or will not rely on an Advice of Counsel defense.

[Meta Response: We said that our response reflected our position and that we would take your issue back to the client. To resolve this issue, Meta will supplement its response in the manner requested. We trust that this issue is now resolved.]

Thank you. Please let us know by what date you will supplement your response to this interrogatory.

17. Additional 30(b)(6) Topics: Meta decline to state whether it would object to any of the amended 30(b)(6) topics Plaintiffs served on October 8, and if so to meet and confer now and attempt to resolve any such objections, because the response time set forth in the rules has not yet elapsed.

[Meta Response: As we told you, we were evaluating your topics and identifying the witnesses that Meta plans to designate. This must be informed by your list of intended 30(b)(1) deponents, which Plaintiffs provided only on Thursday, October 17th (and subsequently have modified over the weekend). Meta will meet and confer with Plaintiffs' this week about the 30(b)(6) topics and any objections.]

Thank you, please let us know when Meta is available to meet and confer this week. Also, please let us know if Meta will identify its designated 30(b)(6) witnesses (and by topic) this week. Absent Meta's agreement to provide this information, we remain at an impasse over this issue.

From: Lauter, Judd <jlauter@cooley.com>

Sent: Monday, October 21, 2024 12:05 AM

To: Reed Forbush <rforbush@BSFLLP.com>; Kathleen Hartnett <khartnett@cooley.com>; Jesse Panuccio <jpanuccio@BSFLLP.com>; Maxwell Pritt <mpritt@BSFLLP.com>; Holden Benon <hbenon@saverilawfirm.com>; Christopher Young <cyoung@saverilawfirm.com>; Aaron Cera <aCera@saverilawfirm.com>; Cadio Zirpoli <czirpoli@saverilawfirm.com>; Joe Saveri <jsaveri@saverilawfirm.com>; Margaux Poueymirou <mpoueymirou@saverilawfirm.com>; Ashleigh Jensen <ajensen@saverilawfirm.com>; Rya Fishman <rfishman@saverilawfirm.com>; Matthew Butterick <mb@buttericklaw.com>; Nada Djordjevic <ndjordjevic@dicellolevitt.com>; James Ulwick <Julwick@dicellolevitt.com>; Bryan L. Clobes <BClobes@caffertyclobes.com>; Mohammed Rathur <MRathur@caffertyclobes.com>; Amy Keller <akeller@dicellolevitt.com>; David Straite <dstraite@dicellolevitt.com>; Ruby Ponce <rponce@saverilawfirm.com>; Alexander Sweatman <ASweatman@caffertyclobes.com>; Heaven Haile <hhaile@saverilawfirm.com>; Llama BSF <Llama_BSF@bsfllp.com>; Josh Schiller <JiSchiller@BSFLLP.com>; David Boies <DBoies@BSFLLP.com>

Cc: Ghazarian, Colette A <cghazarian@cooley.com>; Poppell, Cole A <CPoppell@cooley.com>; Dunning, Angela L. <adunning@cgsh.com>; Ghajar, Bobby A. <bghajar@cooley.com>; Alvarez, Jessica <jalvarezlopez@cooley.com>; Weinstein, Mark <mweinstein@cooley.com>; Stameshkin, Liz <lstameshkin@cooley.com>; z/Meta-Kadrey <zmetakadrey@cooley.com>

Subject: RE: Kadrey v. Meta - October 16 Meet and Confer

CAUTION: External email. Please do not respond to or click on links/attachments unless you recognize the sender.

Counsel:

This email responds to your emails below from Friday and this morning regarding the parties' 10/16 meet and confer, on which the parties discussed Plaintiffs' two 10/9 letters regarding their claimed disputes with Meta's discovery responses.

In this morning's email, you demanded we respond by 9 AM tomorrow to your voluminous set of inquiries and stated that you would be contacting Judge Hixson's chambers "because you did not get back to us on any of the issues despite our request that you do so by Friday." However, we did not agree to provide responses by Friday, but instead told you on our meet and confer that we might not be able to respond until at least Monday. Your Friday email also does not accurately reflect a number of our positions and aspects of the parties' conversation on 10/16, and below we correct the record. We are further troubled by your assertion in this morning's email of additional new issues about which you are demanding our response by tomorrow at 9 AM. This violates both L.R. 37 and Judge Hixson's Discovery Standing Order. We are available to meet and confer with you on the new issues you have raised.

Without waiving any objections and reserving all rights, below is an update regarding the items listed in your Friday email, with corrections to the record as necessary. We have copied the text of your statements from your Friday email below, and we have placed our responses in brackets. Your Friday email divided the issues into two unnamed categories, and for clarity we have renamed the first “Follow-Up Issues” and the second “Potentially Disputed Issues”.

I. **“Follow-Up Issues”**

1. RFP No. 64: Meta will consider proposing an agreeable limitation and searching/producing responsive documents.

[Meta response: This request seeks documents showing “each instance within the last three years where You have licensed copyrighted works for Meta’s commercial use,” without any limitation whatsoever, including any tie to Meta’s training of the Meta Language Models. As we discussed on the call, Meta objected to this request 6 months ago and Plaintiffs never raised it since, and it is unduly burdensome and not proportional to the needs of the case. In response to your latest set of RFPs, Meta will be making a production regarding the licensing of copyrighted textual works for Meta’s use in training the Meta Language Models, which reflects the scope of potentially relevant materials in this case. We believe that this production will be sufficient and appropriate to respond to RFP No. 64 and the latest requests. We trust that this issue is now resolved.]

2. LibGen Training Data (RFP Nos. 1-3 & 7, ROG No. 1): Meta declined to produce all datasets with copyrighted material that it downloaded but will consider whether it would identify all iterations of downloaded datasets with copyrighted material (including the Bashlykov download of LibGen).

[Meta Response: Since the beginning of this case, Meta has been clear that it would produce the training data alleged to contain the Plaintiffs’ works and has done so, along with significant additional training data. The training data produced by Meta constitutes many terabytes of data, and Plaintiffs’ counsel has complained that this volume of data is too large for them to handle. In any event, Meta had understood that the parties had reached a compromise to limit production of datasets to those datasets that are alleged to contain Plaintiffs’ works. It is not proportional to the needs of the case, nor pertinent to adjudication of the fair use defense, to produce copies of all pre-training and fine tuning datasets, but Meta is willing to consider producing additional datasets, if any, that contain books. In addition, Meta will confirm whether the Libgen dataset that has been produced is the current and most comprehensive version of that dataset. However, Meta will not identify and produce duplicative copies of each training dataset, which, among other things, would impose a highly disproportionate burden in light of Plaintiffs’ needs.]

3. GitHub-Hosted Communications and Source-Code Repositories: Meta will investigate gaps in data (e.g., time periods during which there were no pull requests) as identified in declaration of Jonathan Krein.

[Meta Response: We confirm that we are actively identifying any alleged gaps in the pull requests and will produce any located pull requests for the requested time periods. We are also investigating whether additional source code should be produced in response to Plaintiffs’ most recent set of discovery requests (as Plaintiffs’ initial requests did not call for source code). We thus trust that this issue is now resolved.]

4. Llama Filters or Training Data (RFP Nos. 45, 64, 74, 76, and 77): Meta will consider providing screenshot(s) of the filtering tool that Meta's counsel described on the call and represented was not collectible.

[Meta Response: None of the cited RFPs address any alleged tools for filtering data. Moreover, the "tool" you are referring to is merely a tool used to visualize data. Because (among other things) there is no identified RFP to which this tool would be responsive, we do not agree there is anything else to provide in response to this request and we trust that this issue is now resolved.]

5. Valuation and Financials: Meta will confirm the scope of data/documents already searched and produced (and whether the documents pertain to Llama 4—see general impasse regarding Llama 4.)

[Meta Response: Meta has already produced financial-related materials responsive to Plaintiffs' document requests. Meta will be producing additional financial-related documents responsive to Plaintiffs' latest set of RFPs in due course to the extent any such documents exist. Meta has not withheld responsive financial related documents on the basis that they relate only to Llama 4. We thus trust that this issue is now resolved.]

6. Identifying Responsiveness of Document Productions: Meta said it'll consider including in its document production letters an indication of the content of the productions and to what RFPs they're responsive.

[Meta Response: We discussed notifying Plaintiffs when a production addresses a production or formatting issue from a prior production. We are not aware of, and you did not provide, any authority supporting an obligation for a party to provide descriptions of the content of its document productions and to tie those productions to specific RFPs. To resolve this issue, going forward we will provide a description of our productions when we make them. We thus trust that this issue is now resolved.]

7. Identification of Search Terms: Meta declined to identify search terms and hit counts unless it was a mutual exchange and said it would get back to us on when it could exchange search terms and hit counts. Meta said it would identify all non-custodial and custodial sources searched but not all sources with potentially relevant information (see Identification of Search Terms entry under Impasse section).

[Meta Response: As we stated, we are agreeable to engaging with Plaintiffs in a mutual exchange of search terms and hit counts as contemplated by the parties' agreed-upon ESI order. (Doc No. 101, para. 7). This disclosure would also identify which non-custodial and custodial sources of information that were searched with these terms. We do not believe there are any other types of data sources with potentially relevant information that were not searched for and/or collected from, nor is that required by the ESI order. We trust that this issue is now resolved.]

8. RFP No. 59: Meta said it would consider proposing a limitation or definition for "fictional works."

[Meta Response: We asked Plaintiffs to explain what they mean by documents about the output of "fictional works." We need to understand what Plaintiffs are seeking here in order to be able to

respond to this request.]

9. Interrogatory No. 3: Meta agreed to supplement its response.

[Meta Response: Yes, Meta will supplement this response. We trust that this issue is now resolved.]

10. Interrogatory No. 4: Meta said it would consider proposing an agreeable construction of this interrogatory.

[Meta Response: We do not know what you mean by “agreeable construction.” We said that we would take another look at the interrogatory and consider supplementation. We will supplement our response. We trust that this issue is now resolved.]

11. Interrogatory No. 5: Meta said it would investigate Plaintiffs’ description of inchoate agreements as negotiations for licenses, deals in process, or deals that fell apart.

[Meta Response: As an initial matter, we do not understand what you mean by “inchoate agreement,” as by definition an agreement is something that is not inchoate. In any event, we will search for and produce any documents about the acquisition or licensing of training data in response to more recently served RFPs. We will also supplement our response to this interrogatory accordingly. We trust that this issue is now resolved.]

12. Interrogatory Nos. 13 and 14: Meta agreed to answer the sub-parts in Rog Nos. 13 and 14.

[Meta Response: While we maintain and do not waive any objections to the fact that these interrogatories and others previously served by Plaintiffs contain multiple subparts, Meta will supplement its responses. We trust that this issue is now resolved.]

13. Data Formats: Meta confirmed some garbled chats are from WhatsApp and represented its production of Workplace Chats are in the format available. Meta said it would consider providing file types (but did not agree to re-produce chats in any other formats, see entry in Impasses section).

[Meta Response: We are working to reproduce the discrete number of WhatsApp chats that had a processing issue. With respect to Workplace chats, you have not identified any legitimate issues with the format of those materials. All text is legible, including emojis and links, in the document you cited. To your question, the format in which they were archived, searched, and collected has a file extension of eMail. This reproduction and information about file extension should address any issues that exist re “data format.” We trust that this issue is now resolved.]

II. **“Potentially Disputed Issues”**

1. Production of Incorrectly Imaged Documents: Meta said it was aiming for the end of October to re-produce documents incorrectly imaged and/or formatted that Judge Hixson ordered on 10/1 it must produce. Plaintiffs will ask Judge Hixson to impose an earlier deadline given the deposition schedule.

[Meta Response: Meta has already addressed and reproduced 1,335 documents with imaging issues,

which is a majority of the documents at issue. There are no more than 1,100 documents remaining with potential imaging issues. We anticipate completing our review and reproduction by no later than the end of the month and will endeavor to do so sooner. We trust that this issue is now resolved.]

2. General Scope of Custodial Files Searched:

Meta declined to search custodial files outside of the 10 custodians it originally selected and the additional five ordered by Judge Hixson. Meta also represented its search for responsive documents was limited to non-custodial sources and custodial sources of only those 10 custodians (and now the additional five). Meta also identified two for which it searched custodial files beyond its initial 10 custodians but declined to identify whether it conducted a similar search for any other requests. This dispute concerns all RFPs, including Communications with Business Partners and Communications with Shadow Libraries.

[Meta Response: In compliance with the parties' agreed-upon ESI Order, Meta conducted a comprehensive search for documents responsive to Plaintiffs' requests and conducted custodial documents for the designated custodians. The ESI Order limited custodial ESI searches to only 10 custodians, and Judge Hixson's October 4 order added 5 more custodians. You have identified no authority for your demand that Meta conduct boundless custodial searches beyond the 15 custodians that have already been identified. On our meet and confer, we discussed specific requests regarding, for example, training data discussions with certain companies (RFP 29-34) or archives of information on the internet (RFP Nos. 7-12). We are willing to conduct some reasonable follow up searches for these discrete categories.]

Meta declined to add any additional custodians identified by Plaintiffs.

[Meta Response: We have already produced documents for the 10 custodians authorized by the ESI order and are in the process of doing so for 5 additional custodians allowed for by Judge Hixson's recent order. Plaintiffs have identified no reason for doubling (and now in your latest email, tripling) the number of custodians this late in the discovery process. Plaintiffs also have failed to follow the required procedures set forth in the ESI order for seeking additional custodians.]

3. All Llama Models and All Datasets: Meta represented that despite its written objections it has included all Llama models in its document and data searches and productions except as to Llama 4, and it declined to search for and produce responsive information and data in connection with Llama 4. Meta also represented that despite its written objections it has not limited its search for and production of datasets to Book3, but it declined to identify all datasets with copyrighted works that it copied or how many times those datasets, in full or in part, were downloaded. Meta also represented it believed only datasets relating to training was relevant and declined to produce other datasets. This dispute concerns Plaintiffs' RFP Nos. 64, 67, and 49, GitHub-Hosted Communications and Source-Code Repositories, documents and interrogatories related to Valuation and Financials.

[Meta Response: This issue relates to Follow Up Item #2 above, and we refer you to our response above. Meta has stated that it would produce the training data alleged to contain the Plaintiffs' works and has done so. That training data constitutes multiple terabytes of data, which Plaintiffs' counsel has complained is too large for them to handle. It is not proportional to the needs of the case, nor pertinent to adjudication of the fair use defense, to produce any duplicative copies of enormous data sets. With respect to Llama 4, that model is being developed and not expected to be

released until next year, after the close of fact discovery in this case. While we are willing to work with Plaintiff to identify reasonable discovery to address the fair use issues as it relates to Llama 4, we do not think it is proportional to the needs of the case nor practical to comprehensively address another, unreleased and unfinalized model in the short time period the parties have to conduct the remaining discovery.]

4. RFP No. 45: Meta declined to expand its search beyond use of “tools” for identifying licensed material to include documents or materials regarding licensing and communications with book publishers.

[Meta Response: While we maintain our position that there are no “tools” responsive to your request, as discussed above regarding Follow Up Item #1, Meta will be producing documents about acquisition or licensing of training data in response to more recently served RFPs. We consider this issue resolved.]

5. Production of Hyperlinks: Meta declined to make a supplemental production of all hyperlinked documents.

[Meta Response: As the ESI order makes clear, hyperlinked documents are not akin to attachments (ESI Order, Doc. No. 101 at 17). Meta has been responding to your requests for additional hyperlinked documents; has been producing the hyperlinked documents when they are responsive; and has noted that in many instances you already had the document you were asking us to produce. We expect to complete our production of any remaining responsive hyperlinked documents that you have requested in about a week. We trust that this issue is now resolved.]

6. Addition of Timothy Dettmers as Custodian: Meta declined to add Mr. Dettmers as a custodian.

[Meta Response: We do not agree that Mr. Dettmers should be added as a custodian. To begin, the ESI order has a protocol for adding custodians and Plaintiffs have not followed that protocol here. Moreover, as we have repeatedly explained, and as Plaintiffs are well aware, Mr. Dettmers was a part-time employee at Meta prior to the company’s development of the Llama models and was uninvolved in the development of those models. Mr. Dettmers’ documents, including those that Judge Hixson determined were privileged, have no bearing on the issues in dispute. Moreover, you have known about Mr. Dettmers since at least December 2023 and provide no reason why Plaintiffs’ suddenly need his ESI now, particularly after Judge Hixson’s adjudication of the privilege dispute.]

7. Identification of Repositories Not Searched: Meta declined to identify repositories of potentially relevant documents that were not searched. This dispute is relevant to all of RFP Nos. 1-12 and 36-38.

[Meta Response: We do not understand what you are referring to. Meta searched for and/or collected responsive documents in the data sources within Meta’s possession, custody or control that would have potentially responsive or relevant documents, including both non-custodial files and custodial files (see response to Potentially Disputed Issue #2 above). Meta is not aware of any types of data sources with potentially responsive or relevant documents that were not searched. We will provide that information when the parties exchange ESI search terms and hit counts (see Follow Up Issue #7 above).]

8. Identification of Noncustodial and Custodial Sources: Meta declined to identify all sources that potentially might have responsive information even if not searched, and confirmed its investigation of custodial files to search consisted of asking the custodians in a general way what modes of communication they use to communicate about work (without specifically asking about alternative modes like WhatsApp, Discord, Signal, Telegram, etc.). Meta was not able to explain the production of Al-Dahle WhatsApp messages on the night before his deposition.

[Meta Response: Your statements do not reflect what the parties discussed. Meta searched for and/or collected responsive documents in the data sources within Meta's possession, custody or control that would have potentially responsive or relevant documents, including both non-custodial files and custodial files (see response to Potentially Disputed Issue #2 above). For each of the ESI custodians, Meta has made a reasonable inquiry into which modes of communication they use for work, including by asking about particular modes by name. We will identify the data sources (including communication platforms) that were searched when the parties mutually exchange search terms and hit counts.

9. Searching Manifold, GTT, Hive: Meta generally refused to perform searches in Manifold, GTT or Hive on the purported grounds that purportedly they cannot be searched and, in any event, would only contain duplicative data. Meta agreed to consider identifying (not producing) the iterations of copies of training datasets with copyrighted material or books within their PCC.

[Meta Response: This item involves storage locations that may include training data, an issue that was previously addressed in Follow Up Item #2 above. Meta has stated that it would produce the training data alleged to contain the Plaintiffs' works and has done so. That training data constitutes multiple terabytes of data, which Plaintiffs' counsel has complained is too large for them to handle. It is not proportional to the needs of the case, nor pertinent to adjudication of the fair use defense, to produce any duplicative copies of enormous data sets.]

10. RFP No. 54: Meta declined to search for documents on the grounds that the "factual premise" of the RFP is purportedly wrong, even if there had been a decision at one point not to market Llama with a user interface.

[Meta Response: RFP No. 54 asks for documents about a decision "not" to create an interface for users for the Meta Language Models. But Meta did create an interface for users for the Meta Language Models – Meta AI – and Meta has produced documents about that interface. This document request therefore is based on an inaccurate premise and, in light of that inaccuracy, you have not articulated any reason it would be relevant or proportional to Plaintiffs' needs.]

11. Interrogatory No. 15: Meta declined to answer all sub-parts of this interrogatory.

[Meta Response: Reserving all rights and objections, Meta intends to supplement its response to this interrogatory. We trust that this issue is now resolved.]

12. Additional Interrogatories Requested: Meta declined to agree to Plaintiffs' request to expand the numbers of interrogatories the parties may serve and answer the interrogatories Plaintiffs

served on October 10.

[Meta Response: Meta is willing to agree to a small increase in the number of interrogatories, up to 5 more than the 25 the parties had previously agreed to, provided the increase is mutual for both sides. We trust with this compromise that this issue is now resolved.]

13. Data Formats: Meta declined to reproduce in a legible format the Workplace Chat and WhatsApp messages that Plaintiffs explained contain garbled text or was not produced in a reasonably usable format.

[Meta Response: This complaint is duplicative of Follow Up Item #13 above. We have addressed these issues there.]

14. Relevant Time Period: Meta declined to extend its searches to the beginning of the class period.

[Meta Response: Since the beginning of discovery in this case, Meta has clearly stated that its collection and production of documents would be based on the January 1, 2022 date that it included in its responses and objections. This is a conservative date because Meta did not even begin work on the Meta Language Models at issue in this case until late 2022. Plaintiffs never raised any issue with this time frame until your October 9 letter. We have no reason to believe that there would be relevant responsive documents at an earlier date than 2022 regarding the Llama models.

15. Fair Use Interrogatory (No. 1, Set 2): Meta declined to amend its interrogatory to remove its reference to unidentified documents that it contends support its Fair Use defense and to identify the them.

[Meta Response: Your statement does not reflect the parties' discussion. We told you that Meta will supplement its response to this interrogatory, including identifying documents where appropriate. We trust that this issue is now resolved.]

16. Advice of Counsel Interrogatory (No. 2, Set 2): Meta would not agree to amend its response to state without ambiguity and equivocation whether it will or will not rely on an Advice of Counsel defense.

[Meta Response: We said that our response reflected our position and that we would take your issue back to the client. To resolve this issue, Meta will supplement its response in the manner requested. We trust that this issue is now resolved.]

17. Additional 30(b)(6) Topics: Meta decline to state whether it would object to any of the amended 30(b)(6) topics Plaintiffs served on October 8, and if so to meet and confer now and attempt to resolve any such objections, because the response time set forth in the rules has not yet elapsed.

[Meta Response: As we told you, we were evaluating your topics and identifying the witnesses that Meta plans to designate. This must be informed by your list of intended 30(b)(1) deponents, which Plaintiffs provided only on Thursday, October 17th (and subsequently have modified over the

weekend). Meta will meet and confer with Plaintiffs' this week about the 30(b)(6) topics and any objections.]

Regards,
Judd

Judd Lauter

Cooley LLP
+1 415 693 2915 office
jlauter@cooley.com

From: Reed Forbush <rforbush@BSFLLP.com>

Sent: Sunday, October 20, 2024 8:55 AM

To: Hartnett, Kathleen <khartnett@cooley.com>; Lauter, Judd <jlauter@cooley.com>; Jesse Panuccio <jpanuccio@BSFLLP.com>; Maxwell Pritt <mpritt@BSFLLP.com>; Holden Benon <hbenon@saverilawfirm.com>; Christopher Young <cyoung@saverilawfirm.com>; Aaron Cera <aCera@saverilawfirm.com>; Cadio Zirpoli <czirpoli@saverilawfirm.com>; Joe Saveri <jsaveri@saverilawfirm.com>; Margaux Poueymirou <mpoueymirou@saverilawfirm.com>; Ashleigh Jensen <ajensen@saverilawfirm.com>; Rya Fishman <rfishman@saverilawfirm.com>; Matthew Butterick <mb@buttericklaw.com>; Nada Djordjevic <ndjordjevic@dicellolevitt.com>; James Ulwick <Julwick@dicellolevitt.com>; Bryan L. Clobes <BClobes@caffertyclobes.com>; Mohammed Rathur <MRathur@caffertyclobes.com>; Amy Keller <akeller@dicellolevitt.com>; David Straite <dstraite@dicellolevitt.com>; Ruby Ponce <rponce@saverilawfirm.com>; Alexander Sweatman <ASweatman@caffertyclobes.com>; Heaven Haile <hhaile@saverilawfirm.com>; Llama BSF <Llama_BSF@bsfllp.com>; Josh Schiller <JSchiller@BSFLLP.com>; David Boies <DBoies@BSFLLP.com>

Cc: Ghazarian, Colette A <cghazarian@cooley.com>; Poppell, Cole A <CPoppell@cooley.com>; Dunning, Angela L. <adunning@cgsh.com>; Ghajar, Bobby A. <bghajar@cooley.com>; Alvarez, Jessica <jalvarezlopez@cooley.com>; Weinstein, Mark <mweinstein@cooley.com>; Stameshkin, Liz <lstameshkin@cooley.com>; z/Meta-Kadrey <zmetakadrey@cooley.com>

Subject: RE: Kadrey v. Meta - October 16 Meet and Confer

Importance: High

[External]

Good Morning Counsel – Following up on our discovery letters dated 10/8, M&C on 10/16, and the email we sent on Friday.

- A. We will contact Judge Hixson's chambers tomorrow (Monday, 10/21) morning to request a telephonic conference with the parties to address the issues that are highlighted yellow in my email below, per the standing order. We are proceeding on all outstanding issues because you did not get back to us on any of the issues despite our request that you do so by Friday. Please let us know by 9 am pacific tomorrow if you do not think that any highlighted issues can be resolved by telephonic conference so we can state your position when we request the conference.
- B. Please also let us know if you will agree to shorten the time for Meta to provide any

objections to our recent written discovery to 21 days. Plaintiffs likewise would agree to provide any objections and responses within 21 days to the discovery you served Friday night. If not, that is an additional issue we would like to present to Judge Hixson at the conference.

- C. In light of Meta's position that it won't be representing former employee Edouard Grave, we request that he be added as a document custodian. We also request that Meta add other requested deponents as custodians (Yann LeCun, Susan Zhang, Sy Choudhury, Eugene Nho, Logan Kerr, Chris Cox, and Thomas Scialom), as well as Manohar Paluri, Jacob Xu, Movarid Metanat, Andrew Bosworth, Nisha Deo, Chris Marra, and Lukas Belcher. Please let us know if you agree to this request in part or in full (along with our outstanding request identified in our M&C correspondence and below to add further document custodians, including other requested deponents and former employees like Guillaume Lample and Xavier Martinet).
- D. Also in light of Meta's position that it will not be representing former employees Edouard Grave, Guillaume Lample, and Xavier Martinet, we will request that Judge Hixson issue an order to French authorities authorizing Plaintiffs to obtain documents from, and to depose, these witnesses.

Again, please let us know your positions by 9am tomorrow (Monday, 10/21). Thank you.

Best,
Reed
646-339-8787

Reed D. Forbush

Counsel

BOIES SCHILLER FLEXNER LLP

44 Montgomery Street
41st Floor
San Francisco, CA 94104
(t) +1 415 293 6815
(m) +1 646 339 8787
rforbush@bsfllp.com
www.bsfllp.com

From: Reed Forbush <rforbush@BSFLLP.com>

Sent: Friday, October 18, 2024 2:26 PM

To: Kathleen Hartnett <khartnett@cooley.com>; Lauter, Judd <jlauter@cooley.com>; Jesse Panuccio <jpanuccio@BSFLLP.com>; Maxwell Pritt <mpritt@BSFLLP.com>; Holden Benon <hbenon@saverilawfirm.com>; Christopher Young <cyoung@saverilawfirm.com>; Aaron Cera <aCera@saverilawfirm.com>; Cadio Zirpoli <czirpoli@saverilawfirm.com>; Joe Saveri <jsaveri@saverilawfirm.com>; Margaux Poueymirou <mpoueymirou@saverilawfirm.com>; Ashleigh Jensen <ajensen@saverilawfirm.com>; Rya Fishman <rfishman@saverilawfirm.com>; Matthew Butterick <mb@buttericklaw.com>; Nada Djordjevic <ndjordjevic@dicellolevitt.com>; James Ulwick <Julwick@dicellolevitt.com>; Bryan L. Clobes <BClobes@caffertyclobes.com>; Mohammed Rathur <MRathur@caffertyclobes.com>; Amy Keller <akeller@dicellolevitt.com>; David Straite

<dstraite@dicellolevitt.com>; Ruby Ponce <rponce@saverilawfirm.com>; Alexander Sweatman <ASweatman@caffertyclobes.com>; Heaven Haile <hhaile@saverilawfirm.com>; Llama BSF <Llama_BSF@bsfllp.com>; Josh Schiller <JiSchiller@BSFLLP.com>; David Boies <DBoies@BSFLLP.com>

Cc: Ghazarian, Colette A <cghazarian@cooley.com>; Poppell, Cole A <CPoppell@cooley.com>; Dunning, Angela L. <adunning@cgsh.com>; Ghajar, Bobby A. <bghajar@cooley.com>; Alvarez, Jessica <jalvarezlopez@cooley.com>; Weinstein, Mark <mweinstein@cooley.com>; Stameshkin, Liz <lstameshkin@cooley.com>; z/Meta-Kadrey <zmetakadrey@cooley.com>

Subject: Kadrey v. Meta - October 16 Meet and Confer

Dear Counsel,

This correspondence memorializes the outcome of the parties' discovery meet and confer on October 16, 2024, to address Plaintiffs' two discovery deficiency letters to Meta dated and served on October 9, 2024.

The issues you stated that Meta would investigate or otherwise look into in the hope of resolving them are as follows:

1. **RFP No. 64:** Meta will consider proposing an agreeable limitation and searching/producing responsive documents.
2. **LibGen Training Data (RFP Nos. 1-3 & 7, ROG No. 1):** Meta declined to produce all datasets with copyrighted material that it downloaded but will consider whether it would identify all iterations of downloaded datasets with copyrighted material (including the Bashlykov download of LibGen).
3. **GitHub-Hosted Communications and Source-Code Repositories:** Meta will investigate gaps in data (e.g., time periods during which there were no pull requests) as identified in declaration of Jonathan Krein.
4. **Llama Filters or Training Data (RFP Nos. 45, 64, 74, 76, and 77):** Meta will consider providing screenshot(s) of the filtering tool that Meta's counsel described on the call and represented was not collectible.
5. **Valuation and Financials:** Meta will confirm the scope of data/documents already searched and produced (and whether the documents pertain to Llama 4—see general impasse regarding Llama 4.)
6. **Identifying Responsiveness of Document Productions:** Meta said it'll consider including in its document production letters an indication of the content of the productions and to what RFPs they're responsive.
7. **Identification of Search Terms:** Meta declined to identify search terms and hit counts unless it was a mutual exchange and said it would get back to us on when it could exchange search terms and hit counts. Meta said it would identify all non-custodial and custodial sources searched but not all sources with potentially relevant information (see Identification of Search Terms entry under Impasse section).
8. **RFP No. 59:** Meta said it would consider proposing a limitation or definition for "fictional

works.”

9. Interrogatory No. 3: Meta agreed to supplement its response.
10. Interrogatory No. 4: Meta said it would consider proposing an agreeable construction of this interrogatory.
11. Interrogatory No. 5: Meta said it would investigate Plaintiffs’ description of inchoate agreements as negotiations for licenses, deals in process, or deals that fell apart.
12. Interrogatory Nos. 13 and 14: Meta agreed to answer the sub-parts in Rog Nos. 13 and 14.
13. Data Formats: Meta confirmed some garbled chats are from WhatsApp and represented its production of Workplace Chats are in the format available. Meta said it would consider providing file types (but did not agree to re-produce chats in any other formats, see entry in Impasses section).

We asked Meta to get back to us on these issues by today, October 18, and you said you would attempt to do so.

The issues on which the parties are at an impasse and we will proceed to brief for filing on 10/23 are as follows:

1. Production of Incorrectly Imaged Documents: Meta said it was aiming for the end of October to re-produce documents incorrectly imaged and/or formatted that Judge Hixson ordered on 10/1 it must produce. Plaintiffs will ask Judge Hixson to impose an earlier deadline given the deposition schedule.
2. General Scope of Custodial Files Searched:
 - a. Meta declined to search custodial files outside of the 10 custodians it originally selected and the additional five ordered by Judge Hixson. Meta also represented its search for responsive documents was limited to non-custodial sources and custodial sources of only those 10 custodians (and now the additional five). Meta also identified two for which it searched custodial files beyond its initial 10 custodians but declined to identify whether it conducted a similar search for any other requests. This dispute concerns all RFPs, including Communications with Business Partners and Communications with Shadow Libraries.
 - b. Meta declined to add any additional custodians identified by Plaintiffs.
3. All Llama Models and All Datasets: Meta represented that despite its written objections it has included all Llama models in its document and data searches and productions except as to Llama 4, and it declined to search for and produce responsive information and data in connection with Llama 4. Meta also represented that despite its written objections it has not limited its search for and production of datasets to Book3, but it declined to identify all datasets with copyrighted works that it copied or how many times those datasets, in full or in part, were downloaded. Meta also represented it believed only datasets relating to training was relevant and declined to produce other datasets. This dispute concerns Plaintiffs’ RFP Nos. 64, 67, and 49, GitHub-Hosted Communications and Source-Code Repositories, documents and interrogatories related to Valuation and Financials.

4. RFP No. 45: Meta declined to expand its search beyond use of “tools” for identifying licensed material to include documents or materials regarding licensing and communications with book publishers.
5. Production of Hyperlinks: Meta declined to make a supplemental production of all hyperlinked documents.
6. Addition of Timothy Dettmers as Custodian: Meta declined to add Mr. Dettmers as a custodian.
7. Identification of Repositories Not Searched: Meta declined to identify repositories of potentially relevant documents that were not searched. This dispute is relevant to all of RFP Nos. 1-12 and 36-38.
8. Identification of Noncustodial and Custodial Sources: Meta declined to identify all sources that potentially might have responsive information even if not searched, and confirmed its investigation of custodial files to search consisted of asking the custodians in a general way what modes of communication they use to communicate about work (without specifically asking about alternative modes like WhatsApp, Discord, Signal, Telegram, etc.). Meta was not able to explain the production of Al-Dahle WhatsApp messages on the night before his deposition.
9. Searching Manifold, GTT, Hive: Meta generally refused to perform searches in Manifold, GTT or Hive on the purported grounds that purportedly they cannot be searched and, in any event, would only contain duplicative data. Meta agreed to consider identifying (not producing) the iterations of copies of training datasets with copyrighted material or books within their PCC.
10. RFP No. 54: Meta declined to search for documents on the grounds that the “factual premise” of the RFP is purportedly wrong, even if there had been a decision at one point not to market Llama with a user interface.
11. Interrogatory No. 15: Meta declined to answer all sub-parts of this interrogatory.
12. Additional Interrogatories Requested: Meta declined to agree to Plaintiffs’ request to expand the numbers of interrogatories the parties may serve and answer the interrogatories Plaintiffs served on October 10.
13. Data Formats: Meta declined to reproduce in a legible format the Workplace Chat and WhatsApp messages that Plaintiffs explained contain garbled text or was not produced in a reasonably usable format.
14. Relevant Time Period: Meta declined to extend its searches to the beginning of the class period.
15. Fair Use Interrogatory (No. 1, Set 2): Meta declined to amend its interrogatory to remove its reference to unidentified documents that it contends support its Fair Use defense and to identify the them.
16. Advice of Counsel Interrogatory (No. 2, Set 2): Meta would not agree to amend its response

to state without ambiguity and equivocation whether it will or will not rely on an Advice of Counsel defense.

17. **Additional 30(b)(6) Topics**: Meta decline to state whether it would object to any of the amended 30(b)(6) topics Plaintiffs served on October 8, and if so to meet and confer now and attempt to resolve any such objections, because the response time set forth in the rules has not yet elapsed.

Best regards,
Reed

Reed D. Forbush
Counsel

BOIES SCHILLER FLEXNER LLP

44 Montgomery Street
41st Floor
San Francisco, CA 94104
(t) +1 415 293 6815
(m) +1 646 339 8787
rforbush@bsflp.com
www.bsflp.com

The information contained in this electronic message is confidential information intended only for the use of the named recipient(s) and may contain information that, among other protections, is the subject of attorney-client privilege, attorney work product or exempt from disclosure under applicable law. If the reader of this electronic message is not the named recipient, or the employee or agent responsible to deliver it to the named recipient, you are hereby notified that any dissemination, distribution, copying or other use of this communication is strictly prohibited and no privilege is waived. If you have received this communication in error, please immediately notify the sender by replying to this electronic message and then deleting this electronic message from your computer. [v.1 08201831BSF]

This email message is for the sole use of the intended recipient(s) and may contain confidential and privileged information. Any unauthorized review, use, disclosure or distribution is prohibited. If you are not the intended recipient, please contact the sender by reply email and destroy all copies of the original message. If you are the intended recipient, please be advised that the content of this message is subject to access, review and disclosure by the sender's Email System Administrator.

The information contained in this electronic message is confidential information intended only for the use of the named recipient(s) and may contain information that, among other protections, is the subject of attorney-client privilege, attorney work product or exempt from disclosure under applicable law. If the reader of this electronic message is not the named recipient, or the employee or agent responsible to deliver it to the named recipient, you are hereby notified that any dissemination, distribution, copying or other use of this communication is strictly prohibited and no privilege is waived. If you have received this communication in error, please immediately notify the sender by replying to this electronic message and then deleting this electronic message from your computer. [v.1 08201831BSF]

This email message is for the sole use of the intended recipient(s) and may contain confidential and privileged information. Any unauthorized review, use, disclosure or distribution is prohibited. If you are not the intended recipient, please contact the sender by reply email and destroy all copies of the original message. If you are the intended recipient, please be advised that the content of this message is subject to access, review and disclosure by the sender's Email System Administrator.